

**Amendments to the Claims:**

This listing of claims will replace all prior versions, and listings of claims in the application:

**Listing of Claims:**

1. (Currently amended) A method for estimating the copy number of a genomic region in an experimental sample comprising:

- (a) isolating nucleic acid from the experimental sample;
- (b) amplifying at least some regions of the nucleic acid;
- (c) labeling the amplified products;
- (d) hybridizing the labeled amplified products to an array to obtain a hybridization

pattern, wherein the array comprises a plurality of genotyping probe sets for a plurality of SNPs comprising autosomal SNPs, wherein a probe set comprises:

- (i) a plurality of perfect match probes to a first allele of a SNP,
- (ii) a plurality of perfect match probes to a second allele of the SNP,
- (iii) a plurality of mismatch probes to the first allele of the SNP, and
- (iv) a plurality of mismatch probes to the second allele of the SNP,

(e) obtaining a measurement for the SNP in the experimental sample wherein the measurement,  $S$ , is the log of the arithmetic average of the intensities of at least two of the perfect match probes to said first allele or at least two of the perfect match probes to said second allele for the SNP in the hybridization pattern normalized to the  $S$  values of all SNPs genotyped in the experimental sample;

(f) obtaining an  $S$  value for the SNP in each of a plurality of reference samples that are matched to the experimental sample in genotype call at the SNP, wherein the  $S$  value is the log of the average of the intensities of at least two perfect match probes in a reference hybridization pattern normalized to the  $S$  values of all SNPs genotyped in that reference sample;

- (g) calculating the mean ~~and the standard deviation~~ for the reference sample  $S$  values using the values obtained in (f);
- (h) obtaining a log intensity difference by subtracting the mean value obtained in (g) from the value of the measurement,  $S$ , obtained in (e); and
- (i) estimating the copy number of the region including the SNP wherein copy number is estimated using:  $Copy\ Number \approx \exp(b + m \times (\tilde{S}_{jg}^c - \hat{\mu}_{jg}))$  wherein  $\tilde{S}_{jg}^c$  is the log of the average of the intensities of the perfect match probes for a SNP  $j$  of genotype  $g$  in an experimental sample  $c$ , normalized to the  $S$  values of all SNPs genotyped in the experimental sample,  $\hat{\mu}_{jg}$  is the ~~average~~ mean of the ~~normalized~~  $S$  values for SNP  $j$  in a plurality of reference samples of genotype  $g$  at SNP  $j$ ,  $b$  is the y-intercept and  $m$  is the slope of a line defined by plotting intensity values from SNPs of known copy number and obtaining from the estimated copy number an estimated copy number alteration and an estimated direction of copy number change.

2. (Previously presented) The method of claim 1 wherein the  $S$  values for all SNPs genotyped in the experimental sample and in each reference sample are normalized so that the mean of the  $S$  values for all the autosomal SNPs in a sample is zero and the variance is 1.

3. (Original) The method of claim 1 further comprising calculating a p-value for the estimated copy number alteration and determining if the p-value is less than a threshold p-value, wherein the estimated direction of copy number change is significant if the p-value is less than the threshold.

4. (Original) The method of claim 2 further comprising calculating a p-value for the estimated copy number alteration and determining if the p-value is less than a threshold p-value, wherein the estimated direction of copy number change is significant if the p-value is less than the threshold.

5. (Original) The method of claim 1 wherein the  $S$  value is calculated using:

$S = \text{Log}(\frac{1}{X} \sum_{i=1}^X PM_i)$  where  $PM_i$  is the intensity of the perfect match cell of probe pair  $i$  and  $X$  is the number of perfect match probes in a set.

6. (Original) The method of claim 5 wherein  $X$  is between 1 and 30.

7. (Original) The method of claim 5 wherein  $X$  is 20.

8. (Canceled)

9. (Canceled)

10. (Previously presented) The method of claim 1 wherein  $b$  is about 0.693 and  $m$  is about 0.895.

11. (Canceled)

12. (Previously presented) The method of claim 1 wherein the experimental sample is a tumor sample.

13. (Original) The method of claim 1 wherein the experimental sample is a mixture of tumor and normal cells.

14. (Original) The method of claim 1 wherein the experimental sample is a sample that is from a non-cancerous sample.

15. (Original) The method of claim 1 wherein the experimental sample is a sample that is suspected of having a chromosomal anomaly selected from the group consisting of a constitutional anomaly, an acquired anomaly, a numerical anomaly, a structural anomaly and mosaicism.

16. (Previously presented) The method of claim 1 wherein at least some of the SNPs of known copy number are SNPs on the X chromosome.

17. (Original) The method of claim 1 wherein each  $S$  value obtained in (f) that is more than 3 standard deviations from the mean of the  $S$  values is excluded from the estimation of mean and standard deviation of the reference distribution calculated in (g).

18. (Previously presented) The method of claim 1 wherein a second estimate of copy number is obtained by comparing the discrimination ratio, DR, of a SNP in an experimental sample with an average DR from that SNP in a plurality of genotype matched reference samples, where the DR for a probe set with 20 PM/MM probe pairs is calculated using:

$$DR = \frac{1}{20} \sum_{i=1}^{20} \left( \frac{PM_i - MM_i}{PM_i + MM_i} \right)$$

where  $PM_i$  is the intensity of the perfect match cell of probe pair  $i$  and  $MM_i$  is the intensity of the mismatch cell of probe pair  $i$ .

19. (Currently amended) A method of identifying a genomic region that is amplified or deleted in an experimental sample comprising:

hybridizing a nucleic acid sample derived from the experimental sample to a genotyping array and measuring hybridization intensities for a plurality of perfect match probes,  $PM_i$ , where  $PM_i$  is the hybridization intensity of the perfect match probe of probe pair  $i$ ;

calculating a value,  $S$ , for each SNP genotyped by the array in the experimental sample using:  $S = \text{Log}(\frac{1}{X} \sum_{i=1}^X PM_i)$  where  $X$  is the number of  $PM$  probes for an individual SNP;

normalizing a plurality of  $S$  values calculated for SNPs genotyped in the experimental sample so that the mean of the  $S$  values is zero and the variance is one to obtain a plurality of normalized  $S$  values for said experimental sample;

obtaining normalized ~~mean~~  $S$  values for each SNP genotyped by the array in a plurality of reference samples and calculating an average of the normalized  $S$  values for each SNP in said plurality of reference samples matched in genotype at that SNP;

estimating copy number of at least one SNP in the experimental sample to obtain an estimated copy number wherein copy number is estimated using:

$\text{Copy Number} \approx \exp(b + m \times (\tilde{S}_{jg}^c - \hat{\mu}_{jg}))$  wherein  $\tilde{S}_{jg}^c$  is the log of the average of the intensities of the perfect match probes for a SNP  $j$  of genotype  $g$  in an experimental sample  $c$ , normalized to the  $S$  values of all SNPs genotyped in the experimental sample,  $\hat{\mu}_{jg}$  is the average of the normalized  $S$  values for SNP  $j$  in a plurality of reference samples of genotype  $g$  at SNP  $j$ ,  $b$  is the y-intercept and  $m$  is the slope of a line defined by plotting intensity values from SNPs of known copy number;

determining a direction of copy number change for the SNP in the experimental sample by comparing the estimated copy number of the SNP in the experimental sample to the ~~normal~~ copy number of the SNP in a sample of known copy number at the SNP; and

measuring a p-value to determine confidence level in the predicted direction of change.

20. (Canceled)

21. (Previously presented) The method of claim 19 where  $b$  is about 0.693 and  $m$  is about 0.895.

22. (Original) The method of claim 19 wherein the nucleic acid sample is derived from the experimental sample using the whole genome sampling assay (WGS).

23. (Withdrawn) A method for determining if the copy number estimates of two or more consecutive SNPs is significant comprising:

identifying two or more contiguous SNPs that either all show an estimated reduction in copy number or all show an estimated increase in copy number relative to a plurality of reference samples;

calculating  $\tilde{z}_{m,n}$  using  $\tilde{z}_{m,n} = \frac{1}{\sqrt{n-m+1}} \sum_{j=m}^n \hat{z}_{jg} \sim N(0,1)$ ;

converting  $\tilde{z}_{m,n}$  to a probability using the standard  $\Phi$  function to obtain a p-value; and,

concluding that the estimates are significant using a p-value threshold.

24. (Withdrawn) A method of identifying at least one region of loss of heterozygosity comprising:

identifying at least one contiguous stretch of homozygous SNP genotype calls in the genome of an experimental sample;

obtaining a probability,  $\hat{P}_i$  of homozygosity for each SNP in the contiguous stretch

wherein  $\hat{P}_i = \frac{\# \text{ of } AA \text{ or } BB \text{ calls on SNP } i}{\text{total } \# \text{ of genotype calls on SNP } i}$ ;

calculating the probability that each of the SNPs in the contiguous stretch is homozygous

by using:  $\hat{P}(\text{SNP } m \text{ to } n \text{ homozygous}) = \prod_{i=m}^n \hat{P}_i$ ; and,

identifying the region containing the SNPs as a region of loss of heterozygosity if

$\hat{P}(\text{SNP } m \text{ to } n \text{ homozygous})$  is less than a p-value threshold.

25. (withdrawn) The method of claim 24 wherein the contiguous stretch is at least 10 SNPs that are genotyped.

26. (Currently amended) A method for estimating the copy number of a region of loss of heterozygosity comprising identifying a region of loss of heterozygosity in an experimental sample by:

(i) identifying at least one contiguous stretch of homozygous SNP genotype calls in the genome of the experimental sample;

(ii) obtaining a probability,  $\hat{P}_i$  of homozygosity for each SNP in the contiguous

stretch wherein  $\hat{P}_i = \frac{\# \text{ of } AA \text{ or } BB \text{ calls on SNP } i}{\text{total } \# \text{ of genotype calls on SNP } i}$ , wherein the SNP has a first allele,

A, and a second allele, B, and wherein AA is a homozygous call for the A allele and BB is a homozygous call for the B allele;

(iii) calculating the probability that each of the SNPs in the contiguous stretch is

homozygous by using:  $\hat{P}(\text{SNP } m \text{ to } n \text{ homozygous}) = \prod_{i=m}^n \hat{P}_i$ ; and,

(iv) identifying the region containing the SNPs as a region of loss of

heterozygosity if  $\hat{P}(\text{SNP } m \text{ to } n \text{ homozygous})$  is less than a p-value threshold; and

estimating the copy number of a region identified as a region of loss of heterozygosity in

(iv) by a method comprising:

calculating ~~[[an]]~~ a normalized  $S$  value for at least one SNP in the identified region of

loss of heterozygosity in the experimental sample using:  $S = \text{Log}(\frac{1}{X} \sum_{i=1}^X PM_i)$  where  $PM_i$  is the

intensity of the perfect match cell of probe pair  $i$  and  $X$  is the number of probe pairs in a set and

normalizing the  $S$  value;

calculating normalized  $S$  values for the at least one SNP from a plurality of matched  
genotype call reference samples and calculating an average of the reference sample normalized  $S$   
values for the SNP;

comparing the normalized  $S$  value for the SNP in the experimental sample with the  
average of the reference sample normalized  $S$  values for the SNP ~~in the reference sample~~ to  
obtain a ratio; and

estimating copy number of the SNP in the experimental sample using the ratio.

27. (Original) The method of claim 26 wherein copy number is estimated for 2 or more  
contiguous SNPs in the region.

28. (Canceled)



29. (Original) The method of claim 26 wherein the plurality of matched genotype reference samples comprises at least 10 samples.

30. (Currently amended) A computer software product comprising:

computer program code for inputting a plurality of perfect match intensity values ( $PM_i$ ) for each of a plurality of SNPs in each of a plurality of samples comprising an experimental sample and [[or]] a plurality of reference samples sample;

computer program code for calculating an experimental  $S$  value for each SNP in the experimental sample, wherein said experimental  $S$  value is the log of the mean of the perfect match intensity values for each individual that SNP in the experimental each sample, wherein there is a plurality of reference samples;

computer program code for calculating a normalized experimental  $S$  value for each experimental  $S$  value, by normalizing mean each experimental  $S$  value calculated to all of the experimental  $S$  values within individual experimental and reference samples;

computer program code for calculating a reference  $S$  value for each SNP in each of a plurality of reference samples, wherein said reference  $S$  value is the log of the mean of the perfect match intensity values for that SNP in that reference sample;

computer program code for calculating a normalized reference  $S$  value for each reference  $S$  value by normalizing that reference  $S$  value to all of the reference  $S$  values for SNPs in that reference sample to generate normalized  $S$  values for each of the reference samples;

computer program code for calculating an average of the normalized reference  $S$  values ~~a log of the mean of the intensity value~~ for each individual SNP in all reference samples of matched genotype call at that individual SNP;

computer program code for calculating a ~~log intensity~~ difference between the ~~log mean intensity~~ normalized experimental  $S$  value of a SNP from an experimental sample and the mean

of the normalized reference  $S$  values for that SNP in a plurality of reference samples having a genotype matched to the experimental sample at that SNP ~~log mean intensity of that SNP from reference samples matched to the experimental sample in genotype call at the SNP;~~

computer program code for estimating the copy number of the SNP using:

*Copy Number*  $\approx \exp(b + m \times (\tilde{S}_{jg}^c - \hat{\mu}_{jg}))$  wherein  $\tilde{S}_{jg}^c$  is the normalized experimental  $S$  value ~~log of the average of the intensities of the perfect match probes~~ for a SNP  $j$  of genotype  $g$  in an experimental sample  $c$ , ~~normalized to the  $S$  values of all SNPs genotyped in the experimental sample,~~  $\hat{\mu}_{jg}$  is the ~~average~~ mean of the normalized reference  $S$  values for SNP  $j$  in a plurality of reference samples of genotype  $g$  at SNP  $j$ ,  $b$  is the y-intercept and  $m$  is the slope of a line defined by plotting ~~intensity  $S$~~  values from SNPs of known copy number;

computer program code for calculating a p-value for a direction of change wherein said direction of change is determined by comparing the ~~normal~~ copy number of the SNP in a normal reference sample to the estimated copy number of the SNP;

computer program code for determining if the calculated p-value is less than a selected threshold value; and

computer readable media for storing said computer program codes.

31. (Original) The computer software product of claim 30 wherein the log of the mean intensity value for each SNP is calculated using  $S = \text{Log}(\frac{1}{X} \sum_{i=1}^X PM_i)$  where  $X$  is the number of PM probes per SNP.

32. (Canceled)

33. (Canceled)

34. (withdrawn) A computer software product for identifying at least one region of loss of heterozygosity comprising:

computer program code for identifying at least one contiguous stretch of homozygous SNP genotype calls in the genome of an experimental sample;

computer program code for obtaining a probability,  $\hat{P}_i$  of homozygosity for each SNP in the contiguous stretch wherein  $\hat{P}_i = \frac{\text{\# of AA or BB calls on SNP } i}{\text{total \# of genotype calls on SNP } i}$ ;

computer program code for calculating the probability that each of the SNPs in the contiguous stretch is homozygous by using:  $\hat{P}(\text{SNP m to n homozygous}) = \prod_{i=m}^n \hat{P}_i$ ;

computer program code for identifying the region containing the SNPs as a region of loss of heterozygosity if  $\hat{P}(\text{SNP m to n homozygous})$  is less than a p-value threshold; and

a computer readable media for storing said computer program codes.

Claims 35-39 (canceled)